Supplementary online material for Kwon *et al.*

Addendum to the main text

(i) *ATE1*$^{-/-}$ embryos from *ATE1*$^{+/-}$ intercrosses were present at the expected (~25%) frequency up to ~E13.5, but virtually no *ATE1*$^{-/-}$ embryos were recovered alive by E17. Specifically, no *ATE1*$^{-/-}$ mice were recovered amongst either 954 F_2-generation pups of the C57BL/6J-129SvEv (mixed) background or 267 F_2-generation pups of the 129SvEv (inbred) background. Timed intercrosses of *ATE1*$^{+/-}$ mice were used to determine that *ATE1*$^{-/-}$ embryos were present at approximately the expected (25%) frequency up to ~E13.5, but no *ATE1*$^{-/-}$ embryos were recovered alive by E17. Until E12.5, *ATE1*$^{-/-}$ embryos appeared to be morphologically normal; however, their growth stopped during E13.5-E15.5. By E14.5-E15.5, ~50% of *ATE1*$^{-/-}$ embryos were still alive, but growth-retarded. Live E14.5-E15.5 embryos were capable of opening their mouths and flexing their bodies, suggesting the absence of gross neuromuscular defects. Sections through E13.5 *ATE1*$^{-/-}$ embryos indicated the presence and apparently normal appearance of major organs, except for the phenotypes described in the main text and below.

(ii) We examined the expression of *ATE1* mRNA during embryogenesis using Northern hybridization with total RNA from +/+ embryos ranging in age from E4.5 to E18.5. *ATE1* mRNA was present at least as early as E4.5, and a strong spike of *ATE1* expression was observed during E7.5-9.5 (Fig. S1E). The ~2 kb *ATE1* transcript detected in adult mouse testis (*1*) was also clearly present during the spike of *ATE1* expression in E7.5-E9.5 embryos (Fig. S1E). The *ATE1*$^{-}$ allele was marked with NLS-β-galactosidase (hereafter βgal), expressed from the *ATE1* promoter

(Fig. S1A). During E9.5-E12.5, the expression of *ATE1* (βgal) was high in the neural tube, including the floor plate, motor neurons, and the neural fold; *ATE1* was also expressed in dorsal root ganglia, sympathetic ganglia and notochord, in neurons that regulate blood vessels, in sharply delineated subsets of myotomal cells in each somite, in the mesonephric vesicles (kidney precursor), in the gut, in specific areas of the eye primordium, at the tips of limb buds, in the endodermal layer of the yolk sac (but not in its mesodermal layer), and in specific areas of embryonic heart, including trabeculae, atrial septum, endocardial cushion, aortic valve, and aorta (Fig. S2). We also used whole-mount *in situ* hybridization to examine the expression of *ATE1* in E10.5 +/+ embryos. Higher levels of *ATE1* mRNA expression, as detected through *in situ* hybridization, were observed in subsets of myotomal cells in each somite, in the surface ectoderm of limb bud, in the eye primordium, and in the heart (Fig. S2).

(iii) Formation of the two ventricular chambers involves proliferation and maturation of myocytes in the compact muscular layer, accompanied by trabeculation inside the chamber. The septum is formed through condensation of trabeculae at the interventricular groove, with the medial walls of expanding ventricles fusing together, growing inward, and forming the muscular portion of septum (*2*). The defects of *ATE1*$^{-/-}$ hearts described in the main text suggest that the loss of R-transferase activity inhibits proliferation of myocardial cells. In addition, the PTA defect occurs when too few of the neural crest-derived cells populate the cardiac outflow tract, resulting in a failure to separate the common truncal outflow vessel into the aorta and pulmonary artery (*3*). Since *ATE1* is highly expressed in structures produced by cells originating in the neural crest (Fig. 1 of the main text), the PTA defect of *ATE1*$^{-/-}$ hearts suggests that the R-transferase activity may also be required for migration and/or differentiation of neural crest-derived myocardial cells.

(iv) Angiogenic defects in, for example, the yolk sac of an *ATE1*$^{-/-}$ embryo (Figs. 2B, F and S4B, D) might stem, at least in part, from the (expected) pumping defects of abnormal *ATE1*$^{-/-}$ heart. To address this possibility, we compared cardiac defects of *ATE1*$^{-/-}$ embryos and their angiogenic defects in yolk sacs. E13.5-E14.5 embryos were individually scored for angiogenic defects, such as thin primary vessels and poorly organized small vessels, before determining their genotype. The hearts of *ATE1*$^{-/-}$ embryos thus identified were then examined by serial sectioning and staining. Amongst 13 extensively examined *ATE1*$^{-/-}$ embryos, only one had an apparently normal heart (Fig. S4E-H); yet, similarly to other *ATE1*$^{-/-}$ embryos, it exhibited angiogenesis defects in the yolk sac (Fig. S4B, D), suggesting that abnormal angiogenic remodeling was not caused by the cardiogenic defects of *ATE1*$^{-/-}$ embryos.

(v) To begin a search for mouse genes whose expression is significantly altered in the *ATE1*$^{-/-}$ background, we carried out RT-PCR and Northern analyses with total RNA from +/+ and littermate *ATE1*$^{-/-}$ embryos proper, their yolk sacs, and their hearts, using mouse cDNA probes specific for genes that encode the following proteins: NTAN1, UBR1, UBR2 (components of the N-end rule pathway (*4*)); eHand, ICAM2, GATA6, Nfarc1, dHand, NF1, RXRα, FOG2, GATA4, MEF2c, Neuropilin 1, TEF1, N-myc, RARα, ErbB2 (proteins whose functions include heart development); TIE1, TIE2, FLK1, FLK2, FLT1, FLT4, βH1, GATA1, Il-3R, CD34, VEGF, VEGF-B, VEGF-C, VEGF-D, ANG1, ANG2, ANG3, EfnB1, EfnB2, EfnA1, EphB2 (proteins whose functions include vascular development). No significant differences were detected in the expression of these genes between the *ATE1*$^{-/-}$ and +/+ genetic backgrounds (data not shown),

making it more likely that the absence of N-terminal arginylation impacted a previously undescribed circuit.

Methods

Construction and characterization of ATE1$^{-/-}$ mouse strains

Mouse *ATE1* was isolated using screening, with an *ATE1* cDNA fragment (nt 638-1,491), a BAC library (Genome Systems) from 129/SvJ ES mouse cells (*1*). The exon/intron organization of the first ~20 kb of *ATE1* was determined using exon-specific PCR primers to produce genomic DNA fragments flanked by exons. Details of the targeting vector construction (Fig. S1A) are available on request. The vector was linearized with *Hin*dIII and electroporated into CJ7 embryonic stem (ES) cells, followed by selection and identification (*4, 5*) of the correctly targeted *ATE1*$^{+/-}$ ES cell clones with normal karyotype. Standard techniques (*4-6*) were then used to produce chimeric and (subsequently) *ATE1*$^{-/-}$ mice. Phenotypes of *ATE1*$^{-/-}$ embryos were observed mainly with mice of the C75BL/6J-129SvEv (mixed) background, and confirmed in the 129SvEv (inbred) background. RT-PCR, Southern and Northern analyses, and PCR-mediated genotyping of embryos and pups were performed as described (*4-7*).

R-transferase assay

Ub-X-βgal proteins were purified from *E. coli* carrying pKKUbXβgal plasmids (*8*). The R-transferase reaction (50 μl) contained S105 supernatant (0.5 mg of protein per ml) from either EF cells or whole embryos, prepared as described (*9*), Ub-X-βgal or α-lactalbumin (0.2 mg/ml), *E. coli* tRNA (1 mg/ml), *E. coli* aminoacyl-tRNA synthetases (50 μg/ml), puromycin (0.2 mM), bestatin (0.15 mM), 5 mM MG132 (proteasome inhibitor), 0.4 mM Lys-Ala dipeptide (inhibitor

of post-arginylation steps in the N-end rule pathway), 1 mM ATP, 10 mM creatine phosphate, 0.1 M KCl, 5 mM $MgCl_2$, 50 mM β-mercaptoethanol, 50 mM Tris-HCl (pH 8.0) and 0.3 mM ^{3}H-arginine (New England Nuclear). The reaction mixture was incubated for 3 hr (2 hr with embryo extracts) at 37ºC. A 20-µl sample was precipitated with 10% TCA, and analyzed by SDS-12% PAGE and fluorography.

Pulse-chase assays

Primary mouse EFs were established from E13.5 *ATE1*$^{-/-}$ and littermate +/+ embryos as described (*5*), and immortalized (*4*) to increase transfection efficiency. Cells were transiently transfected with pcDNA3flagDHFRhaUBXnsP4flag, which encoded X-nsP4^f, a set of 69K flag-tagged Sindbis virus RNA polymerase proteins bearing different N-terminal residues (*4*). X-nsP4^f proteins were expressed from the P_{CMV} promoter as parts of ^{f}DHFR^h-Ub^{R48}-X-nsP4^f fusions, the constructs of the UPR (Ub/protein/reference) technique (*10-13*). ^{f}DHFR^h-Ub^{R48}-X-nsP4^f is cotranslationally cleaved by deubiquitylating enzymes at the Ub^{R48}-X junction, yielding the long-lived, epitope-tagged ^{f}DHFR^h-Ub^{R48} reference protein and a test protein X-nsP4^f (*4*). Superscripts "f" and "h" denote, respectively, the flag and ha epitopes (*6*). Cells were labeled with ^{35}S-EXPRESS (New England Nuclear) for 10 min at 37ºC, followed by a chase for 0, 1, and 2 hr in the presence of cycloheximide, preparation of extracts, precipitation with anti-flag antibody, SDS-10% PAGE, autoradiography, and quantitation using PhosphorImager (*4, 5, 11*). In other pulse-chases, *ATE1*$^{-/-}$ EFs were co-transfected with a plasmid expressing X-nsP4^f (^{f}DHFR^h-Ub^{R48}-X-nsP4^f) and either pCDNA3yATE1, expressing *S. cerevisiae* ATE1, or pCDNA3yATE1C23A (derived from a plasmid (*14*) supplied by

Dr. C. Pickart, Johns Hopkins University), which expressed $ATE1^{C23A}$, bearing Cys→Ala mutation at position 23.

Protein sequencing

Mouse L cells were transiently transfected with pCDNA3RGS4flagHisx6, which expressed mouse RGS4-flag-His_6 from the P_{CMV} promoter and was constructed from the pCDNA3RGS4 plasmid (*15*). Cell extracts were prepared 30 hr later; RGS4-flag-His_6 was purified using Ni-NTA magnetic agarose beads (Qiagen), then treated with 25 mM iodoacetamide in 7 M urea (*6*), followed by SDS-PAGE, the transfer to Immobilon-P membrane, and sequencing by Edman degradation, using the 476A sequencer (Perkin-Elmer). For mass spectrometry, RGS4-flag-His_6 was treated with 90 mM iodoacetic acid in 8 M urea for 50 min at ~20ºC, then cleaved with CNBr in 55% HCOOH under argon atmosphere for 12 hr in the dark, followed by reverse-phase HPLC and on-line, fragmentation-based mass spectrometric sequencing of peptides.

LacZ staining, immunohistochemical staining, apoptosis assay, and in situ hybridization

For histological examination, embryos were fixed in 4% paraformaldehyde, paraffin-embedded, sectioned, and stained with hematoxylin and eosin. For LacZ staining, embryos, tissues or yolk sacs were removed, fixed in 0.25% glutaraldehyde/PBS for 5 min, rinsed twice with PBS, and stained for 6 hr to overnight at 37°C in X-Gal buffer (1.3 mg/ml potassium ferrocyanide, 1 mg/ml potassium ferricyanide, 0.2% Triton X-100, 1 mM $MgCl_2$, and 1 mg/ml X-Gal in PBS (pH 7.2)). LacZ-stained embryos were post-fixed, photographed, and sectioned after embedding in 30% sucrose/PBS or paraffin. Procedures for whole-mount or section staining with anti-PECAM-1 antibody (clone MEC 13.3, Pharmingen) were done as described (*15*), using HRP-conjugated

secondary antibody (Jackson Laboratory). Apoptosis was analyzed by TUNEL labeling with *in situ* cell death detection kit (Boehringer Mannheim) using fluorescein-dUTP. For *in situ* hybridization, 1.2 kb *ATE1* cDNA fragment (nt 638-1,734) was used as a template for synthesizing either antisense- or sense-strand RNA probes labeled with digoxigenin (*6, 16*).

Fig. S1. Construction and characterization of mouse *ATE1*$^{-/-}$ strains. (**A**) Top, a map of the ~20 kb 5'-proximal region of the ~274 kb mouse *ATE1* gene. Middle, the knockout (KO) vector. Bottom, the deletion-disruption *ATE1*$^{-}$ allele, which lacked the *ATE1* exons 1-3. *ATE1*$^{-/-}$ cells that have been produced through the use of this allele were completely devoid of R-transferase activity, in contrast to congenic +/+ cells (Figs. 3A, S5A).) Solid vertical rectangles, solid and wavy horizontal lines denote, respectively, the exons, the mouse DNA and the pBR322 DNA. The directions of transcription of *NLS-lacZ* (NLS-βgal), *neo* (neomycin), and *tk* (thymidine kinase) are indicated by arrows. H, *Hind*III; N, *Nco*I; X, *Xho*I; B, *Bam*HI. (**B**) Southern analysis of *ATE1*, using a 0.57 kb 5'-probe, a 0.45 kb 3'-probe (indicated by black rectangles in A), and *Hin*dIII/*Bam*HI-cut genomic DNA from +/+, *ATE1*$^{+/-}$, and *ATE1*$^{-/-}$ E13.5 embryos. (**C**) Genotyping E13.5 progeny of an *ATE1*$^{+/-}$ intercross. DNA bands corresponding to the wild-type and mutant *ATE1* alleles are indicated on the right. (**D**) Northern analysis. Electrophoresed total RNA from E13.5 embryos was hybridized with cDNA probes specific for the deleted region of *ATE1* cDNA (nt 132-357, Acc. #AF079096) (upper panel), and with β-actin cDNA (lower panel). (**E**) Northern blot analysis with total RNA from different ages of embryos. Probes were cDNA fragments specific for *ATE1* (nt 638-1,734), *UBR1* (nt 116-1,265; AF061555), *UBR2* (nt 2,763-4,923, unpublished) and β-actin. The bottom panel shows ribosomal RNA bands stained with ethidium bromide.

Fig. S2. Expression of *lacZ* (βgal)-marked *ATE1*$^{-}$ allele (*ATE1*^{nlsLacZ}) of the *ATE1* gene during embryogenesis. (**A**) Whole-mount X-Gal staining of an E10.5 *ATE1*$^{+/-}$ embryo. (**B**-**C**) Whole-mount X-Gal staining of an E12.5 +/+ (B) and *ATE1*$^{-/-}$ (C) embryo. At this age (E12.5), *ATE1*$^{-/-}$ embryos seemed largely normal, albeit already growth-retarded. (**D**-**H**) Transverse sections of E9.5 *ATE1*$^{+/-}$ embryo. (D) Caudal trunk section, with prominent βgal expression in the neural tube (n) and a subset of myotome (arrowhead). (E) The lens and surrounding structures of embryonic eye. (F) Mesonephric vesicles (kidney precursor). (G) Gut (g). (H) Floorplate (fl) of the neural tube, notochord (no), and sympathetic ganglion (sg). (**I-L**) *ATE1*$^{-/-}$ (E10.5) embryos were used here, in part because their βgal staining was disproportionately (more than 2-fold) stronger than with *ATE1*$^{+/-}$ embryos, while being essentially the same in its spatial distribution and relative intensities (*17*). (I) High βgal levels in the neural tube, motor neuron (mn), dorsal root ganglion (dg), sympathetic trunks (st) between midline dorsal aorta (do) and subcardinal vein (sv), and in a subset of myotomal cells of each somite (arrowheads). *ATE1* (βgal) expression was also high in specific myotomal cells of each somite (black arrowheads); see (D) for the equivalent spots of βgal expression in a transverse section of an E9.5 *ATE1*$^{+/-}$ embryo). (J-K) Expression of βgal was higher in the right dorsal aorta (a) (J) than in veins (K). (L) Expression of βgal in a ganglion (g) near the primary head vein (v). (**M**) Staining with antibody to smooth muscle α-actin identified the myotomal regions (denoted by dashed white ovals) that included the spots of βgal (ATE1) staining indicated by arrowheads in (D), (I) and (S); these spots corresponded to narrow rings of ATE1 expression in the intact somites. (**N**) βgal expression at the tip of forelimb bud. (**O**) Cross section of E10.5 *ATE1*$^{+/-}$ yolk sac. me, extraembryonic mesoderm; en, extraembryonic endoderm. (**P, Q**) *ATE1*

mRNA, detected by *in situ* hybridization, in E10.5 +/+ embryos. Whole-mount *in situ* hybridization was carried out with antisense (P) or sense (Q) *ATE1* RNA probes (see Methods). (**R**-**U**) Transverse sections of E10.5 embryo shown in (P). (R) The lens. (S) A caudal trunk section showing the expression of *ATE1* mRNA in the myotomal regions (arrowheads); these spots of *ATE1* hybridization corresponded to βgal staining indicated by arrowheads in (D) and (I). (T) The heart. (U) Forelimb bud (compare with panel N). (**V**) Hematoxylin/eosin-stained transverse section of E13.5 $ATE1^{+/-}$ heart that had been whole-mount-stained with X-Gal. AoV, aortic valve; AS, atrial septum; EC, endocardial cushion. Bars: 0.5 mm (A-C and P), 50 μm (D-O, R-S and U), 20 μm (T) and 100 μm (V).

Fig. S3. Perturbation of angiogenesis in $ATE1^{-/-}$ embryos. (**A**) Whole mounts of E15.5 $ATE1^{-/-}$ embryos showing edema (double arrows). (**B**, **C**) Increased apoptosis, detected using the TUNEL assay (*6*), in the edematous region (dorsal area, sagittal sections) of the above $ATE1^{-/-}$ embryo. (**D**, **E**) Cross sections of similarly located areas in yolk sacs of E13.5 +/+ (D) and $ATE1^{-/-}$ (E) embryos that had been whole-mount stained with anti-PECAM-1 antibody, and counter-stained with eosin. Note characteristically smaller collecting vessels in $ATE1^{-/-}$ sacs. (**F**, **G**) Apparently normal intersomitic arteries but poor development of vessels (indicated by arrowheads) that sprout from these arteries in E9.5 $ATE1^{-/-}$ embryos (F versus G). (**H**-**K**) One example of impaired angiogenesis in $ATE1^{-/-}$ embryos was the head's dorsal area in E13.5 +/+ (H, J) and $ATE1^{-/-}$ (I, K) embryos whole-mount stained with anti-PECAM-1 antibody. Dorsal (H, I) and lateral (J, K) views of the heads. This area was a major site of hemorrhages in $ATE1^{-/-}$ embryos (see the main text and its Fig. 1E.). Bars: 2 mm (A), 200 μm (B, C), 100 μm (D, E), 250 μm (F, G), and 1 mm (H-K).

Fig. S4. Apparently normal cardiogenesis and abnormal angiogenesis in an *ATE1*$^{-/-}$ embryo. (**A**-**B**) E15.5 +/+ (A) and *ATE1*$^{-/-}$ (B) embryos, with yolk sacs and ectoplacental cones intact. The exceptional *ATE1*$^{-/-}$ embryo of this example had the heart of apparently normal appearance (**E-H**), in contrast to other 12 extensively examined *ATE1*$^{-/-}$ embryos. In addition to a growth retardation of the E15.5 *ATE1*$^{-/-}$ embryo, typical of *ATE1*$^{-/-}$ embryos after day ~E12 (A versus B), the *ATE1*$^{-/-}$ yolk sac was pale, with thinner and scarcer blood vessels. (**C**-**D**) Similarly located areas of +/+ and *ATE1*$^{-/-}$ yolk sacs of the same embryo as above, stained with anti-PECAM-1 antibody. The obvious difference in the vascular patterns observed (C versus D) is quite similar to that in Fig. 2E, F of the main text, in which these patterns are of *ATE1*$^{-/-}$ embryo with an overtly abnormal heart. (**E**-**H**) Serial transverse sections of the apparently normal heart of the same (exceptional) *ATE1*$^{-/-}$ embryo. No +/+ heart is shown here; for the appearance of +/+ hearts, see Fig. 1I of the main text. Bars: 1 mm (A-B), 100 μm (C-D), and 200 μm (E-H).

Fig. S5. N-terminal arginylation, and oxidation of N-terminal Cys. (**A**) Cell-free assay for mouse R-transferase using ^{3}H-Arg, S105 extracts from +/+ and *ATE1*$^{-/-}$ embryos, and unlabeled X-βgals (X = Cys, Asp, Glu, Met, Arg) as well as α-lactalbumin as test substrates. Arginylated substrates are indicated on the right. An asterisk and bracket indicate arginylated endogenous proteins in the extracts. (**B**) Mouse *ATE1*$^{-/-}$ EF cells were co-transfected with a plasmid expressing X-nsP4^{f} (^{f}DHFR^{h}-Ub^{R48}-X-nsP4^{f}) and a plasmid expressing either wild-type *S. cerevisiae* ATE1 (pCDNA3yATE1) or its enzymatically impaired mutant (pCDNA3yATE1C23A) (*14*). Cells were pulse-labeled for 10 min with ^{35}S-methionine, and chased for 1 and 2 hr. (**C**) Quantitation of

results in (A). For each time point, the ratio of ^{35}S in X-nsP4^{f} (X = Arg, Asp, Glu, Cys) to ^{35}S in the ^{f}DHFR^{h}-Ub^{R48} reference, at the same time point, was plotted as the percentage of this ratio relative to that for Met-nsP4^{f} at time 0 (*4, 10-12*). Curve designations: Met-nsP4^{f} coexpressed with either a vector alone (open green square), or yeast ATE1 (solid green square), or mutant yeast ATE1 (solid black square). Cys-nsP4^{f}: the same except for diamonds instead of squares, and red instead of green. Asp-nsP4^{f}: the same except for circles instead of squares, and blue instead of green. (**D**) The N-terminal sequence of mouse RGS4 deduced from mass spectrometric data in (E). The M_r of residue 2 (Cys-2 in the arginylated RGS4) identified it as cysteic acid (M_r of 151.1 Da, 48.1 Da higher than M_r of Cys). M* denotes homoserine lactone. (**E**) Mass spectrometric analysis of the CNBr-produced 19-residue N-terminal peptide of RGS4. The observed molecular masses of peaks are shown in red. The differences between observed and expected values are shown in blue.

References

1. Y. T. Kwon, A. S. Kashina, A. Varshavsky, *Mol. Cell. Biol.* **19**, 182 (1999).
2. D. Srivastava, *Annu Rev. Physiol.* **63**, 451 (2001).
3. T. L. Creazzo, E. G. Robert, L. Leatherbury, S. J. Conway, M. L. Kirby, *Annu. Rev. Physiol.* **60**, 267 (1998).
4. Y. T. Kwon, Z.-X. Xia, I. V. Davydov, S. H. Lecker, A. Varshavsky, *Mol. Cell. Biol.* **21**, 8007 (2001).
5. Y. T. Kwon *et al.*, *Mol. Cell. Biol.* **20**, 4135 (2000).
6. F. M. Ausubel *et al.*, Eds., *Current Protocols in Molecular Biology* (Wiley-Interscience, New York, 2001).
7. Y. T. Kwon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7898 (1998).
8. D. K. Gonda *et al.*, *J. Biol. Chem.* **264**, 16700 (1989).
9. K. Takao, K. Samejima, *Biol. Pharm. Bull.* **22**, 1007 (1999).
10. F. Lévy, N. Johnsson, T. Rumenapf, A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4907 (1996).
11. T. Suzuki, A. Varshavsky, *EMBO J.* **18**, 6017 (1999).
12. G. C. Turner, A. Varshavsky, *Science* **289**, 2117 (2000).
13. A. Varshavsky, *Methods Enzymol.* **327**, 578 (2000).
14. J. Li, C. M. Pickart, *Biochemistry* **34**, 15829 (1995).

15. I. V. Davydov, A. Varshavsky, *J. Biol. Chem.* **275**, 22931 (2000).
16. R. A. Conlon, J. Rossant, *Development* **116**, 357 (1992).
17. Data not shown.